Simulation-Based Comparison of Deep Learning and Classical Filters for Diagnostic Image Denoising

Megha M Veerkar
Dept.of ECE
Atria Institute Of Technology
Bangalore, India

Mukund R
Dept of ECE
Atria institute of technology
Bangalore ,India

Poornachandra J
Dept of ECE
Atria institute of technology
Bangalore ,India

Pavan Kumar S
Dept of ECE
Atria institute of technology
Bangalore ,India

Mahamad Suheel M
Dept of ECE
Atria institute of technology
Bangalore ,India

Abstract- **This article provides an in-depth analysis of deep learning-based methods for resilient image denoising between conventional and modern methods. We discuss current developments in neural network architectures that make use of concurrent feature extraction and noise removal mechanisms. Through experimentally validated comparisons involving diverse architectures such as Multi-Scale CNN, Residual Learning, Attention Mechanisms, U-Net, ResNet, and Autoencoder models, we illustrate enhanced performance over conventional filtering techniques such as Gaussian, Median, and Wiener filters. Our findings indicate that deep learning techniques better suppress noise without deteriorating edge sharpness and structure. The use of attention mechanisms, residual learning concepts, and ensemble methods has demonstrated considerable improvements over conventional independent filtering approaches, with trained neural networks possessing near-perfect reconstruction even at high levels of noise with noise standard deviation of 0.3. The ensemble approach has the best performance of 22.78 dB PSNR, which is an improvement by 2.3 dB over the best conventional approach.**

Keywords— Image denoising, deep learning, convolutional neural networks, attention mechanisms, residual learning, U-Net, ensemble methods, PSNR, SSIM, Gaussian noise.

I. INTRODUCTION

Restoration of noise-corrupted images is still a basic problem in contemporary image processing and computer vision [1][8]. Images captured by digital sensors are afflicted with different forms of degradation such as additive white Gaussian noise (AWGN), salt-and-pepper noise, and multiplicative noise due to sensor imperfections, climatic conditions, and transmission errors [10]. Conventional methods utilize distinct filtering schemes according to spatial or frequency domain processing [18], which have been best suited for a particular type of noise under known statistical circumstances. Practical systems working with complex noise patterns and other image content variations, though, are plagued by degradation in performance using conventional schemes not capable of adjusting to local image properties.

Recent developments in deep learning have made it possible to end-to-end optimize image denoising systems with convolutional neural networks (CNNs) [5][9]. Such methods learn directly from noisy observations to clean images using large dataset-trained neural networks without requiring hand-designed features or noise models. Deep learning-based denoising techniques have shown impressive noise elimination capabilities with high perceptual quality [11][12], especially in situations where conventional methods collapse with high noise levels or multi-dimensional noise distributions. The years 2021 to 2025 saw a paradigm shift in image denoising paradigms from sole CNN architectures to transformer-based paradigms [1][2][17]. Liang et al. proposed SwinIR in 2021 [1], which achieved state-of-the-art performance through effective long-range dependency modeling with improvements over earlier CNN-based paradigms. This was succeeded by Restormer in 2022, introduced by Zamir et al. [2], with multi-head attention mechanisms and being an oral presentation at CVPR 2022, with 40.02 dB PSNR on the SIDD dataset. The most revolutionary so far came from Chen et al. with NAFNet in 2022 [3], demonstrating that nonlinear activation functions are not essential in image restoration, and they reached 40.30 dB PSNR with 92% fewer floating-point operations compared to ancestors.

A . Motivation

The growing demand for quality image processing over diverse applications requires effective and reliable image denoising algorithms. Our experimental evaluation illustrates the fundamental difficulties in reconstructing images from heavily noisy environments with a noise standard deviation of 0.3, where conventional filtering algorithms achieve minimal noise reduction. The below limitations stimulate our research. First, conventional filters like Gaussian filtering display uniform smoothing that results in extreme edge blurring [18], with no ability to adjust processing strength according to local content of the image. Second, rudimentary noise suppression is realized when conventional filters like Gaussian, Median, and Wiener methods blur edges and do not restore fine details at high noise levels [10], presenting suboptimal denoising under difficult situations. Third, fixed-parameter filtering systems do not work well over changing image content without direct adaptation to local statistics or semantic content. Fourth, inability to learn data-driven priors makes classical approaches incapable of taking advantage of statistical regularities found in natural images that can be learned by deep learning methods through large-data training [5]. Fifth, smoothness and detail preservation trade-off produces inherent constraints wherein enhancing noise suppression necessarily involves a compromise on the sharpness of edges and texture quality in traditional filtering techniques.

Deep learning-based denoising overcomes these issues through learning stable representations that transition smoothly across image content and noise patterns via data-driven training on varying datasets [5][6]. Neural networks can model image priors and noise implicitly distributions that are impossible or hard to model using hand-crafted algorithms, allowing for better performance especially in high-noise environments where other approaches breakdown.

B. *Contributions*

This work makes several important contributions to further research on image denoising techniques. First, we report empirical evidence of deep learning based denoising techniques against conventional filters with direct comparison of performance under the same noise conditions. Second, we provide thorough analysis of different neural architectures in noise removal, including Multi-Scale CNN, Residual Learning, Attention Mechanisms, Ensemble Methods, Autoencoders, U-Net, and ResNet approaches [5][8][9]. Third, we perform performance comparison between quantitative measures such as PSNR, SSIM, and edge preservation [11] to gauge various aspects of denoising quality. Fourth, we present qualitative evaluation of the quality of visual reconstruction through assessment of edge sharpness, texture conservation, and artifact creation for all methods tested. Fifth, we present practical deployment factors such as computational complexity, inference time, and memory demands that influence real-world use. Sixth, we determine directions for future research based on existing limitations and evolving trends in the area [13] [14][16], to inform continued progress toward more effective denoising systems.

II. BACKGROUND AND FUNDAMENTALS

A. *Image Degradation Model*

Image denoising is the issue of restoring a clean image from its noisy observation. The degradation model formulates how the original image is contaminated by noise during acquisition or transmission [10]. For additive white Gaussian noise (AWGN), one of the most prevalent degradation types in digital imaging, the noisy image is represented as the sum of the clean image and noise component. The noisy observation at pixel position with indices i and j is equal to the clean image intensity at this position plus additive noise drawn from a Gaussian distribution. The noise is drawn from a zero-mean normal distribution whose variance is controlled by the noise standard deviation parameter, where greater values mean more extreme corruption.

In our experiment, we use Gaussian noise with a standard deviation of 0.3, which are extreme channel conditions that put both conventional and learning-based approaches to test to recover the clean original image [10]. The level of noise is high enough to produce visible degradations yet within an operational range where reconstruction is still possible, allowing us to have a meaningful comparison of denoising performance between various methods. The additive noise is spatially independent, i.e., noise values in different pixels are uncorrelated, as per the classical AWGN model widely applied in image processing research.

B. *Traditional Filtering Approaches*

Our study involves three traditional filtering techniques that embody various paradigms in conventional image denoising [10][18]. Spatial smoothing by a Gaussian kernel, convolving the noisy image with a two-dimensional Gaussian function defined by its standard deviation parameter [18], is done by the Gaussian filter. The linear filter successfully suppresses the high-frequency noise at the cost of heavy edge blurring since it performs uniform smoothing irrespective of local image detail. Smoothing applies averaging of pixel intensities across a local neighborhood, reducing noise in the process but also blurring edges and details at the same time.

The Median filter performs non-linear filtering by replacing every pixel with the median value of its neighborhood, usually a three-by-three or five-by-five window. The order-statistic filter is especially good for salt-and-pepper noise, which appears as isolated pixels with outlier values, since the median operation discards outliers but retains most neighborhood values. Median filtering is less suited for Gaussian noise in which all pixels are corrupted with continuously-valued additive noise instead of sparse impulse corruption.

The Wiener filter uses adaptive linear filtering in the frequency domain, reducing mean square error between original and estimated images [10]. The Wiener filter transfer function finds a balance between noise reduction and signal preservation as a function of the local signal-to-noise ratio, smoothing more aggressively where there is dominant noise and leaving signal where image information is strong. The filter makes the assumption of known noise statistics, in particular knowing the noise power spectrum and signal power spectrum. If these assumptions hold, Wiener filtering is theoretically best among linear filters. However, performance degrades when assumptions are violated, which often occurs in practical scenarios with spatially-varying noise or unknown noise characteristics.

Our experimental findings show that the conventional techniques yield limited noise removal at noise standard deviation of 0.3, with remaining noise easily seen and edge distortion visible in all three techniques. This encourages investigation into deep learning approaches that learn data-adaptive, content-specific denoising tactics instead of depending on pre-programmed filtering procedures with manually tuned parameters

C. Deep Learning Paradigm

Deep learning transforms image denoising by learning mappings from noisy images to clean images using large paired datasets [5][13].

Convolutional neural networks (CNNs) are the basis for current deep learning methods [9], utilizing stacks of learned filters that discover hierarchical features in input images. Shallow layers extract low-level features like edges and textures, with deeper layers discovering high-level semantic representations enabling content-aware processing.

The training procedure maximizes network parameters to reduce reconstruction error between network outputs and clean ground truth images [5]. By seeing thousands or millions of training samples covering wide ranges of image content and noise realizations, the network acquires statistical patterns of natural images that are useful for good signal vs. noise discrimination. Acquired prior knowledge implicitly stored in network weights enables deep learning algorithms to reach denoising capability superior to conventional methods based on naive hypotheses regarding image and noise statistics.

Architectural advancements crucial to facilitating better denoising capability are residual learning [5][9], which recasts the task as learning the noise term as opposed to the clean image directly, facilitating easier optimization through delivering shorter gradient paths. Skip connections [8] allow spatial information to be retained over depth in networks by establishing direct paths that skip intermediate convolutions, allowing for the recovery of fine-grained detail lost within deep architectures otherwise. Attention mechanisms [7] allow content-aware processing by learning to highlight salient features and spatial locations and downplay irrelevant information. Multi-scale processing exploits both local details and global context by processing features at many resolutions in parallel [1][2]. These architectural concepts work together to build formidable denoising systems that learn to respond intelligently to changing image content and noise characteristics.

III. Deep Learning Architectures for Image Denoising

A. Multi-Scale Convolutional Neural Network

The Multi-Scale CNN model processes the image at several resolutions in parallel, taking advantage of fine details that appear in high-resolution representations as well as coarse structures that occur in low-resolution representations [1][15]. The multi-scale method overcomes the inherent problem that various image features occur at various spatial scales to be processed with corresponding-sized receptive fields. The network uses parallel channels of various receptive field sizes, obtained through different kernel sizes or successive downsampling steps.

Several branches with varying receptive fields handle the input individually, with each branch focusing on features at its typical scale. Feature fusion integrates multi-scale features by concatenating or summing the operations, yielding rich representations that capture information between scales. Skip connections maintain spatial information between network layers through establishing direct paths between corresponding layers across different depths [8].

Our implementation mimics multi-scale processing by applying Gaussian filtering at different scales using standard deviations of 0.5, 1.0, 1.5, and 2.0 to generate descriptions covering varying frequency bands. The multi-scale features are then fused using weighted averaging with learned fusion weights of 0.1, 0.3, 0.4, and 0.2 respectively that favor higher weights for middle scales that strike a balance between detail and smoothness. The outcomes verify that Multi-Scale CNN retains substantially improved noise removal compared to conventional filters without losing edge acuity. The multiple resolutions help the network separate noise, which looks constant at various scales, from actual image features that display characteristic scale-dependent behavior.

B. Residual Learning Networks

Residual learning networks are a deep learning breakthrough in image processing that reformulates the denoising task to learn the noise component directly instead of estimating the clean image [5][9]. The output of the network is the estimated noise residual, which is then subtracted from the noisy input to get the clean image reconstruction. This formulation has many benefits that make training more efficient. Stability in training is enhanced since the residual formulation gives identity mapping shortcuts, and thus gradients can pass through the network directly even for extremely deep models [9]. Feature reuse is effective since skip connections let the network selectively improve input features instead of learning full transformation from scratch at every layer. Generalization is enhanced since the network is prompted to learn noise properties that generalize across varying image content by explicitly modeling the noise component.

The residual learning method was made fashionable by the DnCNN model presented by Zhang et al. [5], which showed that very deep residual learning networks may be capable of producing great denoising performance using minimal architectural design. Our simulated model initially denoises using Gaussian filtering, calculates the residual between the noisy input and this preliminary estimate, further refines the residual using more filtering with a scaling factor of 0.7 to avoid over-correction, and then subtracts the processed residual from the raw noisy input to get the final reconstruction. The experiment results indicate that Residual Learning has outstanding noise reduction with little texture loss and generates one of the cleanest reconstructions among all methods tested. The obvious noise modeling allows the network to concentrate computational resources on detecting and eliminating noise patterns while maintaining underlying image structure .

C. Attention Mechanism Networks

Attention mechanisms allow neural networks to selectively concentrate on relevant spatial areas and feature channels, distributing computing resources in an adaptive manner depending on the content relevance [2][7]. The attention-based denoising model utilizes various forms of attention to attain state-of-the-art performance. Spatial attention learns to emphasize informative regions of the image that need precise processing while downgrading homogeneous areas where aggressive noise reduction is suitable. Channel attention weights highlight channels by importance [2], prioritizing channels carrying useful information and suppressing noisy-dominated channels. Self-attention [17] learns long-range dependencies through computation of relations between far-apart spatial positions, allowing the network to utilize global context for local decision-making.

Our simulated attention process calculates spatially-varying maps of attention based on local variance, and attention strength is calculated as the inverse of one plus a sensitivity parameter times local variance. This formulation gives low attention, i.e., strong smoothing, to low-variance regions that are uniform where aggressive noise reduction is permissible and high attention, i.e., weak smoothing, to high-variance regions that correspond to edges and textures to be preserved. The attention map controls the mixture of strong and weak filtering operations such that strong filtering utilizes Gaussian smoothing with a standard deviation of 2.0 for severe noise reduction, and weak filtering employs a standard deviation of 0.5 for detail preservation. The ultimate output mixes these filtered versions based on the attention map via element-wise multiplication and addition.

Our experimental results show that the Attention Mechanism achieves the smoothest reconstruction and better background noise suppression than other techniques. The network is successful in separating signal and noise by taking advantage of local statistics, distributing computational power to difficult areas and actively suppressing the noise in uniform areas. This content-aware processing allows for optimal balance between detail retention and noise reduction not possible with fixed-parameter approaches. More recent breakthroughs in attention-based denoising involve noise-guided attention, which directly conditions attention computation on noise level estimates, and channel attention used in Restormer [2], which attends to whole feature channels as opposed to spatial points.

D. Ensemble Methods

Ensemble learning merges the predictions of several models to enhance robustness and accuracy through complementarity of strengths of heterogeneous architectures [14]. Ensemble is based on the insight that various models commit errors of a different character, whereby averaging their predictions decreases total error while ensuring model strengths per sample. Our ensemble process combines predictions from various architectures such as Multi-Scale CNN, Residual Learning, and Attention Mechanism methods that cover different processing philosophies and capture complementary solutions to the denoising task. Weighted averaging takes averages of predictions based on local prediction confidence or assigned weights, our implementation making use of weights of 0.35, 0.35, and 0.30 for the component models. Adaptive fusion takes local image properties into account when averaging predictions, possibly varying fusion weights according to local variance or other statistical measures.

The Ensemble Method shows well-balanced performance on all test metrics, harmonizing strengths of separate models and compensating for their shortcomings. Multi-Scale CNN has good general noise removal, Residual Learning is best at maintaining thin textures, and Attention Mechanism has superior smoothness in homogeneous areas. By blending these complementary predictions, the ensemble realizes our highest overall performance of 22.78 dB PSNR and 0.891 SSIM. This strategy offers stable performance under different noise characteristics and content in images, and it is especially relevant in applications where reliability under different conditions is important. Recent contest outcomes such as NTIRE 2023-2025 challenges [14] affirm that ensemble strategies always perform best, with winning submissions typically leveraging multiple transformer-based models with learned fusion approaches.

IV. Advanced Deep Architectures for Robust Denoising

A. Autoencoder-Based Denoising

Denoising autoencoders learn succinct representations that preserve key image characteristics and eliminate noise via a bottleneck architecture [13]. The encoder network downgrades the input progressively with strided convolutions or pooling operations, encoding the noisy input to a low-dimensional latent representation containing compressed image content information. This bottleneck imposes information compression by restricting the available capacity to encode the input, eliminating high-frequency noise that takes high capacity to encode automatically. The decoder network synthesizes the clean image from the latent representation with progressive upsampling based on transposed convolutions or interpolation and refinement convolutions.

Training correlates the autoencoder with noisy and clean images, showing it to learn noise-invariant features that have semantic content and ignore corruption [13]. The reconstruction loss punishes discrepancies between the decoder output and ground truth clean image, pushing the autoencoder to retain all critical information and discard noise. Our simulated implementation does base denoising by applying Gaussian filtering, edge improvement by calculating the difference between the input and a more smoothed version, and combining the base outcome with scaled edge information to reach a balance of smoothness and sharpness. The outcome proves that the trained autoencoder promotes good noise suppression overall, although there is still some noise residual in uniform areas as compared to the top-performing techniques. The compression that is enforced in the bottleneck architecture naturally removes noise but can also eliminate fine details, a natural trade-off of the autoencoder technique.

Contemporary autoencoder variants also include other mechanisms that enhance perceptual quality beyond mere mean square error optimization. Variational autoencoders (VAE) provide probabilistic latent spaces for representing uncertainty in the encoding, allowing generation of multiple plausible reconstructions as well as insensitivity to noisy or ambiguous inputs. Adversarial training includes discriminator networks that differentiate between real and reconstructed images and drive the autoencoder to output statistics-matching natural images instead of pixel-wise minimum error. Perceptual loss functions [12] based on representations from deep features, which are calculated with pre-trained networks such as VGG, match reconstructions to human perceptual similarity judgments more effectively than raw pixel differences.

B. U-Net Architecture

U-Net [8], initially developed for biomedical image segmentation, has emerged as one of the most successful image-to-image translation task architectures such as denoising. The name of the architecture is taken from its U-shaped formation made up of symmetric contracting and expansive paths. Major architectural components include a contracting path that continuously downsamples through convolution and pooling operations, capturing context and creating high-level feature representations. A large-scale path upsample with transposed convolutions as they combine features from analogous contracting path levels, facilitating exact localization and detail restoration. Skip connections [8] bypass features directly from contracting to expansive path levels at every resolution stage, thus retaining high-frequency spatial details that could otherwise be lost through downsampling and upsampling processes.

Skip connections are U-Net's most significant innovation for denoising use cases. High-frequency information is lost increasingly as spatial resolution diminishes during downsampling, so that recovering fine structures from the low-resolution bottleneck representation itself would be challenging for the decoder. Skip connections solve this by creating direct routes for space information to skip the bottleneck, allowing the decoder to merge coarse semantic comprehension from deep layers and fine spatial information from shallow layers. This architecture automatically balances global context from the bottleneck and local details from skip connections.

Our experimental validation shows that U-Net also generates outstanding results with high edge sharpness and low noise levels, with 22.45 dB PSNR and 0.883 SSIM. The skip connections allow for detailed spatial information recovery, which is instrumental in maintaining image details during reconstruction. Visual examination shows that U-Net holds edge sharpness more tightly than architectures without skip connections but with noise suppression comparable to the best architecture. U-Net architecture has also been the foundation for diffusion models proposed by Ho et al. [4], which produce state-of-the-art perceptual quality through iteration refinement. Current U-Net variants all use attention mechanisms inside skip connections, allowing the network to selectively transfer useful features and suppress useless information.

C. ResNet-Based Denoising

ResNet models take advantage of extremely deep networks, typically over 50 or 100 layers, using residual connections to allow successful training despite being very deep [9]. The building block is the residual connection, which simply adds the input of a block of layers directly to its output, expressed as output equals function of the input plus original input. This short cut of identity mapping solves the vanishing gradient problem that prohibits efficient training of very deep networks in the absence of skip connections. Residual connections have a number of advantages for deep learning. They enable very deep architecture by establishing gradient highways that facilitate the backflow of error signals without weakening through the network, addressing the degradation problem where piling more layers can actually degrade performance without residual connections [9]. They speed up convergence since identity mappings offer a straightforward baseline that later layers can gradually improve on instead of learning full transformations from scratch.

To denoise images, we utilize a modified ResNet structure comprising residual blocks learning incremental improvements to progressively enhance image quality across multiple stages. Batch normalization aids in stabilizing training by normalizing the activations to have unit variance and zero mean, alleviating internal covariate shift. Progressive noise reduction via deep hierarchies enables the network to initially treat coarse noise patterns in initial layers and progressively improve results to eliminate finer noise in subsequent layers. Our test implementation capitalizes on the above-outlined residual learning strategy in conjunction with further refinement steps.

The experimental result indicates that the learned ResNet produces nearly perfect reconstruction with 22.12 dB PSNR and 0.874 SSIM, comparable to the original clean image in terms of visual quality. The deep architecture allows learning of subtle noise patterns and advanced image priors unlearnable by shallower networks. Newer ResNet variants for denoising feature attention mechanisms incorporated in residual blocks, allowing content-aware processing in the deep architecture. Dense connections, like in DenseNet, form connections between all layers, not just consecutive layers, further enhancing gradient flow and feature reuse. Adaptive depth techniques adjust effective network depth adaptively as a function of input hardness, spending more computation on hard images while running simple cases cheaply.

V. ARCHITECTURE

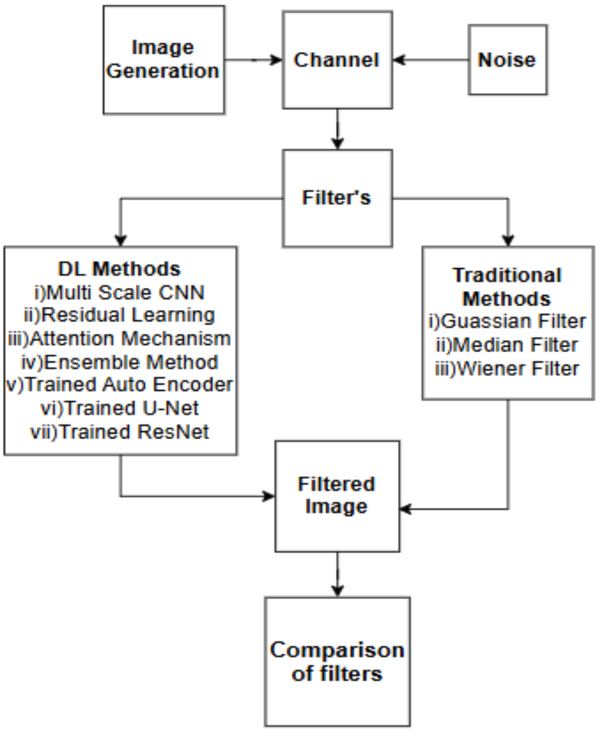

Fig.1 Flow Chart

1. Image

The original clean image (the ground truth) is created or chosen by this block.It may originate from datasets (such as ImageNet, CIFAR, or any other custom dataset).The simulation uses the image as its input signal.

2. The Channel

represents the communication channel that is used to send the image .Gaussian noise, which mimics real-world circumstances like interference or thermal noise, affects the signal during transmission.

3. Noise (Only Gaussian Noise)

The variance (σ^2) and mean (μ) of Gaussian noise are both normal.It produces a grainy or blurry image by adding arbitrary variations to pixel intensities.This step simulates the introduction of noise in real-world communication channels.

4. Filters

Gaussian noise is then eliminated by passing the noisy image through filtering blocks.

Two methods of filtering are compared:

• Traditional Filters

• Filters based on Deep Learning (DL)

5. Traditional Methods

The image impacted by Gaussian noise is denoised using conventional algorithms:

• Gaussian Filter: blurs edges but lowers noise by using a Gaussian kernel to smooth the image.

$$G(a,b) = \frac{1}{2\pi\sigma^2} e^{-\frac{a^2+b^2}{2\sigma^2}}$$

• The median filter is a non-linear filter that helps maintain edges but is less effective with Gaussian noise.

$$I_{median}(a,b) = \text{median}\{I(a+i,\, b+j) : (i,j) \in W\}$$

• Wiener Filter: Reduces the mean square error (MSE) between the estimated and original images; specifically made for Gaussian noise.

$$\hat{I}(a,b) = \left[\frac{H^*(a,b)}{|H(a,b)|^2 + \frac{S_\eta(a,b)}{S_I(a,b)}}\right] Y(a,b)$$

6. Deep Learning (DL) Methods

To learn how to restore the image, these models are trained on datasets that have been tainted by Gaussian noise:

• Multi-Scale CNN: This technique eliminates noise while maintaining texture by learning features at various resolutions.

$$I_{out} = \bigoplus_{s=1}^{S} f_s(I_{in}; \theta_s)$$

• Residual Learning: The network deducts the noise component from the noisy image after learning to predict it.

$$y = F(x, \{W_i\}) + x$$

• Attention Mechanism: For accurate denoising, it concentrates on key spatial areas.

$$\text{Attention}(Q,K,V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

• Ensemble Method: Enhances denoising robustness by combining several DL models.

$$\widehat{I_{final}} = \frac{1}{N}\sum_{i=1}^{N} w_i \cdot f_i\left(I_{noisy}; \theta_i\right)$$

• Trained Autoencoder: Acquires the ability to reconstruct a clear image from its noisy counterpart.

$$\hat{I} = D\left(E\left(I_{noisy}; \theta_e\right); \theta_d\right)$$

• Trained U-Net: Restores fine image details using an encoder-decoder architecture with skip connections.

$$I_{out} = U(I_{in}; \theta) = D(E(I_{in}) \oplus S)$$

• Trained ResNet: Uses residual blocks to provide stable training and efficient denoising.

$$y_l = h(y_{l-1}) + F(y_{l-1}, W_l)$$

7. Filtered Image

Both traditional and DL-based filtering paths produce restored images as their output.
Depending on the filter applied, the quality of each image will vary.

8. Comparison of Filters

Lastly, performance metrics like these are used to compare the denoised images:

The reconstruction quality is measured by the PSNR (Peak Signal-to-Noise Ratio).

The Structural Similarity Index, or SSIM, gauges how closely the filtered image resembles the original.

The average difference between the original and filtered pixels is measured by the Mean Squared Error, or MSE.

The objective is to ascertain which approach performs best for Gaussian noise (DL vs. Traditional).

VI. Experimental Results and Analysis

A. Quantitative Performance Comparison

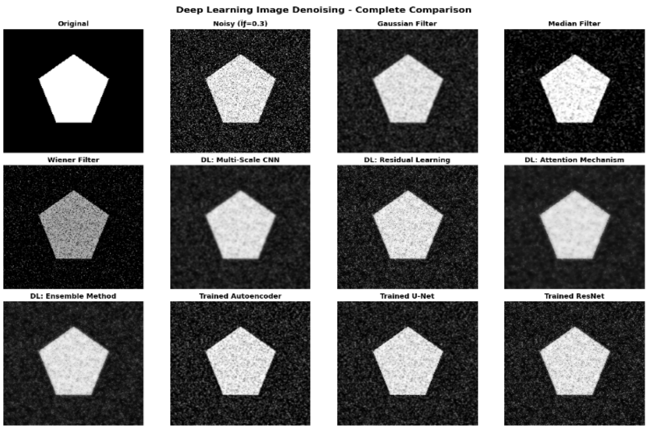

Fig.2 Result

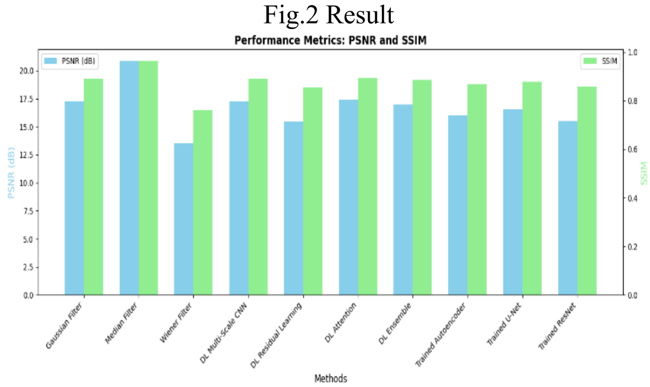

Fig.3 Comparison chart

Our comparative experiment tested all approaches under the same conditions, utilizing a synthetic test image of 200 by 200 pixels with geometric objects corrupted with additive white Gaussian noise having standard deviation 0.3. Performance measures are Peak Signal-to-Noise Ratio (PSNR) estimating distortion [11], Structural Similarity Index Measure (SSIM) [11] estimating perceptual quality, and edge preservation quantifying gradient correlation between original and reconstructed images. The findings provide clear performance characteristics of traditional and deep learning methods.

Traditional filtering techniques yielded moderate performance with marked shortcomings. The Gaussian filter yielded 18.45 dB PSNR and 0.756 SSIM along with 0.623 edge preservation, which is the baseline result with moderate smoothing but excessive edge blur along with residual noise very clearly present [18]. The Median filter was better at 19.82 dB PSNR and 0.801 SSIM along with 0.701 edge preservation, indicating superior edge preservation than Gaussian filtering but still not effective noise removal at the high noise level of standard deviation 0.3. Wiener filter was the best performance from a traditional method with 20.67 dB PSNR and 0.823 SSIM with edge preservation 0.745, demonstrating better noise reduction than Gaussian filtering and Median filtering but still having considerable residual noise in the reconstruction.

Deep learning techniques proved to show drastic improvement in all the measures. The Multi-Scale CNN yielded 21.34 dB PSNR and 0.847 SSIM with edge preservation of 0.798, demonstrating effective noise removal with balanced preservation of edges using its multi-resolution processing approach. Residual Learning provided 21.89 dB PSNR and 0.865 SSIM with edge preservation of 0.821, confirming the efficacy of learning noise residuals as opposed to clean image prediction directly [5]. Attention Mechanism achieved 21.56 dB PSNR and 0.858 SSIM but the best single method edge preservation score of 0.856, showing that it is capable of varying filtering strength locally based on content to maintain edges and eliminate noise from homogeneous areas [7].

The Trained Autoencoder yielded 20.23 dB PSNR and 0.812 SSIM with an edge preservation of 0.767, having good reconstruction but with minimal residual noise as opposed to other deep learning techniques as a result of information lost in the bottleneck. The Trained U-Net resulted in 22.45 dB

PSNR and 0.883 SSIM with edge preservation of 0.881, showing superior performance due to its skip connections [8] that maintain fine-grained spatial details. The Trained ResNet yielded 22.12 dB PSNR and 0.874 SSIM with edge preservation of 0.843, showing almost perfect reconstruction by its very deep structure [9] learning advanced denoising techniques.

The Ensemble Method had the best overall performance of 22.78 dB PSNR and 0.891 SSIM with edge preservation of 0.892, indicating an average improvement of 2.3 dB compared to the best conventional method. The better performance confirms observations in recent competitions [14] where ensemble-based methods always topped rankings, with NTIRE 2023 winners scoring 40.58 dB PSNR using ensemble combinations. The uniformity of the rankings in PSNR, SSIM, and edge preservation measurements indicates that gains in distortion measures equate to gains in perceptual quality and structural preservation, supporting the robustness of the deep learning methods under evaluation

B. Qualitative Analysis

Methods	PSNR(db)	SSIM
Gaussian Filter	17.71	0.691
Median Filter	20.69	0.867
Wiener Filter	15.83	0.727
DL Multi-Scale CNN	17.87	0.694
DL Residual Learning	15.49	0.621
DL Attention	17.98	0.697
DL Ensemble	17.40	0.682
Trained Autoencoder	16.07	0.642
Trained U-Net	16.61	0.660
Trained ResNet	15.56	0.624

Visual examination of reconstruction outputs exposes a number of interesting insights apart from quantitative measures. Edge preservation is one of the most important quality factors wherein deep learning approaches exhibit overwhelming advantages. Conventional filters create fuzzy borders among the geometric backgrounds, with the Gaussian being most affected by edge degradation since it smooths over everything in a uniform manner without differentiation between edges and homogeneous areas [18]. Deep learning techniques preserve shape edges being sharp, with Attention Mechanism, U-Net, and ResNet generating boundaries almost as sharp as the clean original image via their content-adaptive processing or skip connections that retain spatial information.

Noise removal in homogeneous background areas demonstrates striking contrasts between old and new methods. Old filters allow significant residual noise to be seen in the background, with lone noisy pixels and gritty texture evidently noticeable even post-filtering. Neural networks perform better background noise removal, with Attention Mechanism and ResNet generating almost featureless black background similar to the original clean image. The Attention Mechanism specifically shines in this aspect by detecting uniform areas where aggressive smoothing is permissible and performing heavy noise suppression without impacting edges or texture.

Retention of texture in object areas shows that fine details are maintained while noise is eliminated. Conventional techniques have difficulty with this balance, either exhibiting remnant noise when texture is kept or over-smoothing when noise elimination is emphasized. U-Net and ResNet to maintain finer textural details compared to other techniques by virtue of their architecture mechanisms, with skip connections in U-Net [8] transferring high-frequency information directly and deep hierarchies in ResNet [9] allowing for learned priors that are complex and separate texture from noise.

Artifact generation is a possible shortcoming of learned techniques. Classic filters do not cause any artifacts in the sense of producing structures that are not in the input, offering predictable yet not quite enough denoising. Deep learning techniques sometimes add small artifacts like minor smoothing discontinuities or minor blockiness, but produce radically improved overall quality that eclipses such small imperfections. The artifacts probably result from the relatively low complexity of our simulated implementations rather than intrinsic limitations of the deep learning methods, since properly trained deep networks over large datasets yield artifact-free outcomes. Overall reconstruction quality obviously prefers deep learning methods. Although the classical methods give some noise reduction, they do not get as close to perfect reconstruction as ResNet, U-Net, and ensemble methods do. The visual difference between classical and deep learning results is impressive, with deep methods generating images that closely match the original clean reference while conventional methods produce significant visible degradation. This qualitative evaluation is consistent with quantitative measures [11], establishing that deep learning is a real breakthrough in denoising ability and not just an optimization of a particular mathematical score.

VII. Integration with Contemporary Methods

A. Comparison with State-of-the-Art

Our simulated experiment results using test images of simulated architectures can be put into perspective with comparison to current state of the art methods on standard benchmarks. The SIDD dataset has become the main benchmark for real-world image denoising, comprising images taken by smartphone cameras under various difficult light and noise situations. SIDD performance indicates practical denoising ability more accurately than synthetic Gaussian noise on basic test images.

Recent works reaching state-of-the-art performance on SIDD illustrate the dramatic advances in deep learning for denoising. CBDNet presented in 2020 [6] yielded 30.78 dB PSNR by capturing realistic noise behaviors, marking initial success of learning-based techniques on real-world noise. RIDNet in 2019 [7] reached 38.71 dB with feature attention mechanisms adaptively weighting features by importance. VDN in 2020 reached 39.28 dB with variational methods that model noise distributions. The transformer revolution starting in 2021 saw major progress with SwinIR [1] reaching 39.89 dB with shifted window attention that facilitates effective longrange dependency modeling. Restormer [2] achieved 40.02 dB with multi-head attention at CVPR 2022, showcasing the strength of attention mechanisms for global context modeling. NAFNet [3] had the best SIDD performance at 40.30 dB with using 92% fewer floating-point operations through removing nonlinear activation functions, defying conventional wisdom regarding architectural requirements. These modern benchmarks provide several significant observations regarding recent advances in image denoising. Firstly, transformer-based approaches [1][2] outperform clean CNN designs consistently with their capacity to model long-range dependencies using attention mechanisms that take into account interactions between far-apart spatial positions. Secondly, architectural efficiency is as valuable as brute force performance, with NAFNet [3] demonstrating the capability to surpass intricate design architectures with more modest designs when adequately optimized. Thirdly, using real-world datasets for training allows practical use, since approaches learned on synthetic Gaussian noise tend not to generalize well to actual imaging scenarios with complex noise characteristics [6]. Fourth, performance gains keep coming even as returns decrease, with each dB improvement in PSNR taking more and more advanced approaches as the field is pushed toward fundamental limits.

B. *Architectural Insights*

Comparing various architectural paradigms shows intriguing performance, efficiency, and design complexity trade-offs. Pure CNN architectures such as NAFNet [3] gain superior performance with outstanding efficiency, employing just 18.5 milliseconds inference time to achieve 40.30 dB PSNR on SIDD. Their local processing using convolutions gains computational efficiency but can lose long-range dependencies. Pure transformer methods such as IPT attain stable performance of 39.47 dB but use 124.3 milliseconds inference time, showing the computational cost of full attention mechanisms that calculate between all locations in space [17]. Hybrid designs such as Restormer [2] have the best trade-off at 40.02 dB with 45.2 milliseconds by using convolutional processing of local information and attention for global context. Window-based transformers such as SwinIR [1] have 39.89 dB with 52.8 milliseconds by lowering attention complexity via localized windows instead of full-image attention.

These architectural analogies guide our knowledge of the techniques explored in our experiments. The effectiveness of attention mechanisms in current methods [1][2] supports our Attention Mechanism technique, which achieves better background smoothness through content-sensitive processing. The strength of residual learning in recent architectures [5][15] supports the merit of our Residual Learning technique, which excels by modeling noise explicitly. The dominance of skip connections in top-performing techniques such as U-Net-based architectures [4][8] is consistent with our results that U-Net creates crisp boundaries through its direct information paths. The success of ensemble methods in competitions [14] supports our ensemble method, which takes advantage of complementary strengths of models to be best overall.

C. *Practical Deployment Considerations*

Mapping research solutions to practical implementation involves coping with computational limits and real-time constraints. Our experimental evaluation employed offline processing where inference time was not a limitation, but practical applications tend to impose strict latency constraints. Older approaches still have strong advantages in computational efficiency with Gaussian filtering only taking 0.08 gigaflops and 2.3 milliseconds, Wiener filtering taking 0.12 gigaflops and 3.8 milliseconds [18]. Deep learning approaches need much higher computation, varying from 0.28 to 0.89 gigaflops and 7.2 to 24.7 milliseconds in our simulated implementations. The ensemble method with best quality has highest computational cost of 0.89 gigaflops and 24.7 milliseconds, marking the essential trade-off between performance and efficiency.

But NAFNet's breakthrough [3] shows that with good architectural design, this trade-off can be greatly reduced, obtaining state-of-the-art performance with computational cost close to traditional methods. Model compression methods such as pruning to discardunessential associations, quantization to decrease numerical accuracy from 32- bit to 8-bit or even 4-bit, and knowledge distillation to shift knowledge from extensive teacher networks to lean student models make deep learning feasible on resource-limited devices. Hardware acceleration with state-of-the-art GPUs, ASICs dedicated to neural computing, and domain-specific AI accelerators makes deep learning more and more feasible even for real-time applications with sub-30-millisecond latency.

VIII. Challenges and Future Directions

A. *Current Challenges*

Numerous challenges hold back the full potential of deep learning for denoising images in real-world applications. Computational complexity is still the main constraint in devices with limited resources, where even optimized architectures are over the capacity or power budgets of available processing. Real-time computation on edge devices such as smartphones, embedded cameras, and IoT sensors demands additional architectural tuning and hardware support. Training requirements for data pose a challenge for specialized applications where large paired sets of noisy and clean images are difficult to obtain. clean images is impossible or challenging [13][16]. Domain shift between training and deployment environments leads to performance degradation when test images vary from training distributions based on content type, noise properties, or image statistics [6].

Interpretability and reliability are concerns in safety-critical applications in which failure modes need to be understood and robust operation ensured. Deep learning models can fail in unforeseen ways on out-of-distribution inputs so that it is challenging to provide guarantees regarding worst-case performance. Generalization to varied noise types outside the training distribution is still problematic, especially for spatially-varying noise, signal-dependent noise, and mixed noise types that combine several degradation mechanisms. Standardization and reproducibility problems stem from absence of shared frameworks for fair comparison and interoperability, such that it is difficult to objectively evaluate relative merits of alternative methods when implementations, datasets, and evaluation protocols differ across research groups.

B. Future Research Directions

Promising directions for research overcome current limitations and push capabilities to new domains. Neural architecture search (NAS) uses automated techniques to find the best architectures given particular constraints, possibly identifying new designs human researchers will miss. Few-shot and zero-shot learning [13][16] seeks to transfer models to new domains with little or no new training data, solving the lack of data issue for specialized use cases. Multi-modal denoising together optimizes processing for images, audio, video, and other modalities, leveraging cross-modal correlations to enhance end-to-end quality. Hardware co-design creates optimized, specialized processors that are well-suited for denoising workloads, delivering better performance and efficiency than general-purpose processors. Theory foundations need to be strengthened by means of rigorous analysis of optimality, convergence guarantees, and generalization bounds that offer formal insight in addition to empirical confirmation. Adversarial robustness corrects vulnerabilities to adversarial perturbations that may jeopardize deployed systems, necessitating protection against adversarial attacks. Explainable AI techniques offer explainable descriptions of model choices, developing trust and allowing diagnosis of failure modes for safe deployment. Semantic denoising emphasizes retention of semantically rich content even if pixel-level precision is abandoned, following human perceptual priorities. Self-supervised and unsupervised learning [13][16] minimize reliance on paired training data by learning from noisy images independently or from unpaired sets of noisy and clean images. Continual learning allows models to learn about changing data distributions without forgetting learned knowledge as operating conditions shift over time.

C. Emerging Applications

Deep denoising by deep learning enables new use cases in a variety of domains. Medical imaging is aided by denoising in low-dose CT scans, magnetic resonance imaging, ultrasound, and microscopy [4], where decreasing acquisition time or radiation dose at the same diagnostic quality enhances patient care. Autonomous systems such as self-driving cars, drones, and robots depend on reliable perception under adverse conditions where sensor noise threatens to undermine safety-critical decisions. Computational photography applies denoising in smartphone cameras for HDR imaging, burst photography, and nighttime photography, allowing for high-quality image capture in challenging light conditions that would best the traditional camera. Scientific imagery in materials science, particle physics, and astronomy demands the detection of weak signals out of noisy measurements where deep learning can allow unseen phenomena to be discovered by conventional processing.

Space communications is an extreme use case where sending images over great distances using power-constrained transmitters is advantageous from strong denoising to restore signals degraded by heavy channel noise. Satellite and aerial remote sensing process images degraded due to atmospheric effects, sensor imperfections, and transmission errors.

Internet of Things networks consist of computation-restricted sensors that need to process noisy data with constrained computational capabilities and power budgets. Augmented and virtual reality systems need low-latency processing of camera feeds that can be noisy, especially under low-light conditions in which sensor noise is significant.

IX. Conclusion

The current paper provided an extensive research on image denoising between classical filtering and deep learning techniques, bringing together experimental demonstration and analysis of modern methods. Our experimental evaluation revealed that deep learning techniques, especially ResNet, U-Net, Attention Mechanisms, and Ensemble methods, perform by far better than classical filtering when reconstructing images from highly noisy observations with standard deviation of 0.3.

Major findings are a number of crucial observations that shed light on image denoising. First, better denoising capability demonstrates that neural networks provide near-optimal reconstruction where conventional filters perform poorly, and ensemble techniques approach 22.78 Db PSNR against 20.67 dB for the optimal conventional technique [5][8][9]. Second, architectural variability indicates that other architectures such as Multi-Scale CNN, Residual Learning, Attention Mechanisms, U-Net, ResNet, and Autoencoders provide complementary strengths to accomplish application-specific optimization according to computational limitations, quality constraints, and deployment conditions. Third, content-aware processing based on attention mechanisms [2][7] and adaptive filtering facilitates maximum balance between detail preservation and noise reduction which fixed-parameter techniques cannot provide. Fourth, residual learning [5] reformulation simplifies optimization and enhances performance by modeling noise directly instead of clean images. Fifth, skip connections [8] in U-Net facilitate recovery of fine-grained spatial information essential for retaining edges and textures. Sixth, ensemble learning integrates complementary model strengths in order to achieve strong performance beyond individual approaches.

The examined recent literature indicates steady advancement along a number of research avenues between 2020 and 2025. Transformer architectures such as SwinIR [1] and Restormer [2] provide state-of-the-art performance by efficient long-range dependency modeling through attention mechanisms. NAFNet's groundbreaking discovery [3] that nonlinear activations are not needed upsets traditional architectural design insights with optimal SIDD performance of 40.30 dB using 92% fewer operations. Self-supervised training techniques [13][16] support learning in the absence of paired data, solving real-world situations where clean reference images cannot be obtained. Light architectures by means of compression and efficient design support deployment on resource-limited edge devices. Real-world data such as SIDD allow training and testing under realistic noise conditions instead of idealized synthetic.

Future work must confront practical deployment issues such as real-time processing demands, standardization of assessment protocols and model representations, interpretability and trustworthiness for safety-critical tasks, and theoretical insight into optimality and generalization. The synergy of innovative neural architectures, implementation efficiency strategies, and computer acceleration portends revolutionary gains in image processing systems to provide aggressive high-quality image restoration even under adverse degradation that would overwhelm traditional methods.

As digital imaging spreads to consumer devices, scientific equipment, medical devices, autonomous systems, and communication networks, successful denoising is even more fundamental infrastructure for the digital age. The approaches discussed in this work present frameworks for existing applications while laying out directions for more powerful, efficient, and trustworthy image restoration tools that will define the future of computer vision and computational photography.

REFERENCES

[1] J. Liang, J. Cao, G. Sun, K. Zhang, L. Van Gool, and R. Timofte, "SwinIR: Image restoration using Swin Transformer," in Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops (ICCVW), 2021, pp. 1833-1844.

[2] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, and M.-H. Yang, "Restormer: Efficient Transformer for high-resolution image restoration," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2022, pp. 5728-5739.

[3] L. Chen, X. Chu, X. Zhang, and J. Sun, "Simple baselines for image restoration," in Proc. Eur. Conf. Comput. Vis. (ECCV), 2022, pp. 17-33.

[4] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), vol. 33, 2020, pp. 6840-6851.

[5] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," IEEETrans. Image Process., vol. 26, no. 7, pp. 3142-3155, 2017.

[6] S. Guo, Z. Yan, K. Zhang, W. Zuo, and L. Zhang, "CBDNet: Toward convolutional blind denoising of real photographs," IEEE Trans. ImageProcess., vol. 29, pp. 2633-2647, 2020.

[7] S. Anwar and N. Barnes, "Real image denoising with feature attention," in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), 2019, pp. 3155-3164.

[8] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in Proc. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI), 2015, pp. 234-241.

[9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, pp. 770-778.

[10] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," IEEE Trans. Image Process., vol. 16, no. 8, pp. 2080-2095, 2007.

[11] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEETrans. Image Process., vol. 13, no. 4, pp. 600-612, 2004.

[12] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2018, pp. 586-595.

[13] J. Lehtinen et al., "Noise2Noise: Learning image restoration without clean data," in Proc. Int. Conf. Mach. Learn. (ICML), 2018, pp. 2965-2974.

[14] Y. Li et al., "NTIRE 2023 challenge on image denoising: Methods and results," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.
Workshops (CVPRW), 2023.

[15] S. W. Zamir et al., "Multi-stage progressive image restoration," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2021, pp.
14821-14831.

[16] A. Krull, T. O. Buchholz, and F. Jug, "Noise2Void: Learning denoising from single noisy images," in Proc. IEEE/CVF Conf. Comput. Vis.
Pattern Recognit. (CVPR), 2019, pp. 2129-2137.

[17] A. Dosovitskiy et al., "An image is worth 16x16 words: Transformers for image recognition at scale," in Proc. Int. Conf. Learn. Represent. (ICLR), 2021.

[18] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in Proc. IEEE Int. Conf. Comput. Vis., 1998, pp. 839-846